Milan Stehlik

Some Properties of Exchange Design Algorithms Under Correlation

Working Paper

Some Properties of Exchange Design Algorithms Under Correlation

Milan Stehlík

Department of Statistics and Mathematics
Wirtschaftsuniversität Wien

Research Report Series

Report 28
February 2006

http://statmath.wu-wien.ac.at/

Some properties of exchange design algorithms under correlation

Milan Stehlík
Department of Statistics and Mathematics,
University of Economics Vienna,
Augasse 2-6, A-1090 Vienna, Austria
e-mail: Milan.Stehlik@isis.wu-wien.ac.at, Stehlik@fmph.uniba.sk

February 9, 2006

Abstract

In this paper we discuss an algorithm for the construction of D-optimal experimental designs for the parameters in a regression model when the errors have a correlation structure. We show that design points can collapse under the presence of some covariance structures and a so called nugget can be employed in a natural way. We also show that the information of equidistant design on covariance parameter is increasing with the number of design points under exponential variogram, however these designs are not D-optimal. Also in higher dimensions the exponential structure without nugget leads to collapsing of the D-optimal design when also parameters of covariance structure are of interest. However, if only trend parameters are of interest, the designs covering uniformly the whole design space are very efficient. For illustration some numerical examples are also included.

Keywords: Design of experiments, Brimkulov algorithm, D-optimality, optimum design, correlation, information matrix, nugget effect, Matérn class of covariance functions, domain and infill asymptotics

1 Introduction

Here we consider the nonlinear model

$$Y(x) = \eta(\theta, x) + \varepsilon(x)$$

with the design points $x_1, ..., x_N$ taken from a compact design space X. The vector parameter $\theta = (\theta_1, ..., \theta_p)^T \in \Theta$ is unknown, and $\eta(.)$ is a known function and the variance-covariance structure $\text{cov}(x_i, x_j, r)$ depends on another unknown vector parameter $r = (r_1, ..., r_q)^T \in B$. Here we consider the Matérn class of covariance functions. The class is motivated by the smoothness of the spectral density, the wide range of behaviors covered and the interpretability of the parameters. This family includes the exponential class (as a special case with $v = 0.5$) and the Gaussian family of correlation function as a limiting case with $v \to \infty$. There are several parametrizations of the Matérn family, here we follow the one given by [Handcock and Wallis 94]

$$\text{cov}(x_i, x_j, \phi, v) = \frac{1}{2^{v-1}\Gamma(v)} (\frac{2\sqrt{v}d}{\phi})^v K_v(\frac{2\sqrt{v}d}{\phi}).$$

Here ϕ and v are the parameters, d is Euclidean distance between points x_i, x_j and K_v is the modified Bessel function of the third kind and order v (see [Abramowitz and Stegun 65]). The parameter v controls the differentiability of the process and ϕ measures how quickly the correlations decay with distance (range parameter in geostatistical literature). Such isotropic covariance model has many applications, for instance in studying the relationship of both yield and quality of wheat to variables of soil property such as temperature and moisture (see Zhu and Zhang, 2005).

The design $x_1, ..., x_N$ is good if it gives precise estimators of the parameters. We have information matrices

$$M_\theta(n) = \frac{\partial \eta^T(\theta)}{\partial \theta} C^{-1}(r) \frac{\partial \eta(\theta)}{\partial \theta^T}$$

and

$$M_r(n) = \frac{1}{2} tr \left\{ C^{-1}(r) \frac{\partial C(r)}{\partial r} C^{-1}(r) \frac{\partial C(r)}{\partial r^T} \right\}.$$

So for both parameters of interest we have

$$M(n, \theta, r) = \begin{pmatrix} M_\theta(n) & 0 \\ 0 & M_r(n) \end{pmatrix}.$$

We can find applications of various criteria of design optimality for second-order spatial models in the literature. Here we discuss D-optimality, which corresponds to the maximization of the determinant of a standard Fisher information matrix.

Theoretical justifications for using the Fisher information in regular normal models with small variances of $Y(s)$ can be found in [Pázman 04]. However there are also some asymptotical justifications.

Currently there are two main asymptotical frameworks, increasing domain asymptotics and infill asymptotics, for obtaining limiting distributions of maximum likelihood estimators of covariance parameters in Gaussian spatial models with or without a nugget effect. These limiting distributions differ in some cases. Zhang and Zimmerman (2005) have investigated the quality of these approximations both theoretically and empirically. They have found, that for certain consistently estimable parameters of exponential covariograms approximations corresponding to these two frameworks perform about equally well. For those parameters that cannot be estimated consistently, however, the infill asymptotics is preferable. In our paper we consider mainly the infill asymptotics. They have also observed, that the Fisher information appears to be a compromise between the infill asymptotic variance and the increasing domain asymptotic variance. For another justification of infill asymptotics see [Abt and Welch 98]. They found, either analytically or by simulation, that the inverse of the Fisher information matrix may well serve as an approximation for the mean squared error matrix of covariance parameter estimators in special cases. They assume that $Y(t)$ were observed at $t_i = (i-1)/(n-1)$ for $i = 1, .., n$ in $X = [0, 1]$, but their results holds also for case when points become dense everywhere in X as $n \to \infty$. They assume a covariance matrix of the form $\sigma^2 R$, where

$(R)_{i,j} = \text{cov}(t_i, t_j)$ can be of the form $1-r|t_i-t_j|$ for $r \in (0,2)$, $\exp(-r|t_i-t_j|)$ (exponential covariance) and $\exp(-r(t_i-t_j)^2)$ (Gaussian covariance) for $r > 0$. [Zhang 04] has showed that not all parameters in a Matérn class are consistently estimable under infill asymptotics. However, we will show in Section 2 that for exponential variogram an equidistant n-point design $\{t, t+d, t+2d, ...\}, n \leq 5$ the covariance parameter information $M_r(n)$ is increasing with n, more exactly we have

$$M_r(n) = (n-1)M_r(2).$$

For exponential variogram some infill asymptotic justification can be found in Zhang and Zimmerman (2005). This provides some theoretical support of the variogram parameters estimation.

Also [Zhu and Stein 04] use the simulations (under Gaussian random field and Matérn covariance) to study whether the inverse Fisher information matrix is a reasonable approximation of the covariance matrix of maximal likelihood estimators as well as a reasonable design criterion. They have observed that when the sample size is small, inverse Fisher information matrices underestimate the variance of ML estimators. As sample size increases, the relative error becomes smaller and smaller. They have already observed that the Fisher information matrix does give good estimate of the variance of ML estimators when the sample size is large enough.

Although some simulation and theoretical studies shows that the inverse Fisher information matrix is not a good approximation of the covariance matrix of the ML estimates it can still be used as a design criterion if the relationship between these two are monotone, since for the purpose of optimal designing the only correct ordering is important. For instance, [Zhu and Stein 04] observes a monotone relationship between them.

From now on, we mean by information the Fisher information on the studied parameter of the isotropic random field. In this paper we discuss the structure of Fisher information matrices $M_\theta(n)$ and $M_r(n)$ under the exponential covariance structure. We also study an algorithm for the construction of D-optimal experimental designs for the parameters. We show that design points can collapse under the presence of some covariance structures. Here a so called *nugget* can be employed in a natural way. The numerical examples are also included.

2 Exponential covariance structure.

In this section we are concerning the exponential semivariogram $\gamma(d) = 1 - e^{-rd}$, which is the special case of Matérn semivariogram with zero nugget parameter obtained for $v = 0.5$ and introducing a new range parameter $r = \frac{\sqrt{2}}{\phi}$. Let us consider for simplicity the trend $\eta(\theta, x) = \theta_1$ being constant.

2.1 $M_\theta(n)$ structure

Without the loss of generality let the design space be $X = [-1, 1]$. Proposition 2 in [Stehlík 04] states, that $\{-1, 1\}$ is a D-optimal two-point design. If we consider three-point-design, $\{t_1, t_2, t_3\}$, $-1 \leq t_1 < t_2 < t_3 \leq 1$, then Fisher's information $M_\theta(3)$ has form

$$1 + \frac{2 + 2\exp(-d_{12} - 2d_{23}) - 2\exp(-d_{12}) + 2\exp(-2d_{12} - d_{23}) - 2\exp(-2(d_{12} + d_{23})) - 2\exp(-d_{23})}{\exp(-2(d_{12} + d_{23})) - \exp(-2d_{12}) - \exp(-2d_{23}) + 1}$$

where $d_{12} = t_2 - t_1$ and $d_{23} = t_3 - t_2$. In [Stehlík 04] is proved, that the design $\{-1, 0, 1\}$ is D-optimal in the exponential setup. Let us suppose 4-point design $-1 \leq t_1 < t_2 < t_3 < t_4 \leq 1$, and denote the distances $d_{12} = t_2 - t_1$, $d_{23} = t_3 - t_2$ and $z = t_4 - t_3$. Then the (intercept) Fisher's information has the form

$$M_\theta(4) = 2(-2 + e^{-d_{34}} + e^{-d_{12}} + e^{-d_{23}} + e^{-2d_{23}-d_{12}-2d_{34}} + e^{-2d_{12}} + e^{-2d_{23}} + e^{-2d_{34}} - e^{-2d_{12}-2d_{23}-2d_{34}}$$

$$-e^{-d_{12}-2d_{34}} + e^{-2d_{12}-2d_{23}-d_{34}} - e^{-d_{34}-2d_{12}} - e^{-d_{23}-2d_{34}} - e^{-2d_{12}-d_{23}} - e^{-2d_{23}-d_{34}} + e^{-2d_{34}-2d_{12}-d_{23}} - e^{-2d_{23}-d_{12}})/$$

$$(-1 + e^{-2d_{34}} + e^{-2d_{23}} - e^{-2d_{23}-2d_{34}} + e^{-2d_{12}} - e^{-2d_{12}-2d_{34}} - e^{-2d_{12}-2d_{23}} + e^{-2d_{12}-2d_{23}-2d_{34}})$$

and the optimum design is the equidistant one with $d_{12} = d_{23} = d_{34} = 2/3$ and the information $M = 1.964538$. To compute this, we can use the direct algorithm which compares the value of the criteria function with the value of the prior design and exchange the compared design with the prior one when the information of the prior one is dominated by compared design.

Due to the knowledge of the analytical form of the information and its properties one can employ Lipschitz and continuous optimization (see [Horst and Tuy 96]), which can be implemented like a net-searching algorithm. The only problem of such an algorithm is its time complexity. The (intercept) Fisher's information in the case of 5-point design $-1 \le x_1 < x_2 < x_3 < x_4 < x_5 \le 1$, and denote the distances d_{12}, d_{23}, d_{34} and d_{45} has much more complicated form

$$M_\theta(5) = (-5 - 2e^{-d_{12}-2d_{23}-2d_{34}-2d_{45}} - 2e^{-2d_{12}-2d_{23}-d_{34}-2d_{45}} + 2e^{-d_{34}-2d_{45}-2d_{12}} - 2e^{-2d_{12}-2d_{34}-2d_{45}-d_{23}} +$$

$$-2e^{-d_{34}-2d_{45}} + 2e^{-2d_{23}-d_{34}-2d_{45}} + 2e^{-2d_{45}-d_{23}-2d_{12}} + 2e^{-2d_{23}-2d_{34}-d_{45}} - 2e^{-2d_{12}-2d_{23}-2d_{34}-d_{45}} + 2e^{-2d_{23}+}$$

$$-d_{12} - 2d_{45} + 2e^{-2d_{34}-d_{45}-2d_{12}} + 3e^{-2d_{12}} + 2e^{-d_{12}} - 2e^{-d_{45}-2d_{23}} + 2e^{-2d_{23}-2d_{12}-d_{45}} + 2e^{-2d_{34}-2d_{45}-d_{23}}$$

$$-2e^{-d_{45}-2d_{12}} - 2e^{-d_{23}-2d_{45}} - e^{-2d_{34}-2d_{45}} - 2e^{-2d_{34}-d_{45}} - e^{-2d_{23}-2d_{34}-2d_{45}} + 3e^{-2d_{45}}$$

$$-e^{-2d_{45}-2d_{23}} + 2e^{-2d_{23}-d_{12}-2d_{34}} + 2e^{-d_{34}} + 2e{-}d_{23} - 2e^{-2d_{23}-d_{12}} - 2e^{-d_{12}-2d_{34}} + 2e^{-2d_{34}-2d_{12}-d_{23}}$$

$$-2e^{-d_{23}-2d_{34}} - 2e^{-2d_{12}-d_{23}} + 3e^{-2d_{34}} + 3e^{-2d_{23}} - 2e^{-2d_{23}-d_{34}} - 2e^{-2d_{45}-d_{12}} - e^{-2d_{23}-2d_{34}} + 2e^{-2d_{12}-2d_{23}-d_{34}}$$

$$+2e^{-d_{12}-2d_{34}-2d_{45}} + 2e^{-d_{45}} - 2e^{-d_{34}-2d_{12}} - e^{-2d_{12}-2d_{34}} - e^{-2d_{12}-2d_{23}} - e^{-2d_{12}-2d_{23}-2d_{34}} + 3e^{-2d_{12}-2d_{23}-2d_{34}-2d_{45}}$$

$$-e^{-2d_{34}-2d_{45}-2d_{12}} - e^{-2d_{23}-2d_{12}-2d_{45}} - e^{-2d_{45}-2d_{12}})/(-1 + e^{-2d_{12}} - e^{-2d_{34}-2d_{45}} + e^{-2d_{23}-2d_{34}-2d_{45}} + e^{-2d_{45}}$$

$$-e^{-2d_{45}-2d_{23}} + e^{-2d_{34}} + e^{-2d_{23}} - e^{-2d_{23}-2d_{34}} - e^{-2d_{12}-2d_{34}} - e^{-2d_{12}-2d_{23}} + e^{-2d_{12}-2d_{23}-2d_{34}} - e^{-2d_{12}-2d_{23}-2d_{34}-2d_{45}}$$

$$+e^{-2d_{34}-2d_{45}-2d_{12}} + e^{-2d_{23}-2d_{12}-2d_{45}} - e^{-2d_{45}-2d_{12}})$$

and computationally we obtain that the optimal design $d_{12} = d_{23} = d_{34} = d_{45} = 1/2$ has information $M = 1.979674635$ (note, that information is increasing with number of design points, what generally does not hold).

2.2 $M_r(n)$ structure

The Fisher information M_r for the covariance parameter r is more complex. Let us study D-optimal designs for (r, θ_1) under the exponential covariance structure. The design space X is some fixed compact subset of R^d with diameter 2. In [Stehlík 04] is shown, that when r not being parameter of interest, the criterion function of the intercept is $M = \frac{2}{1+\exp(-rd)}$, and the D-optimal design $\{-1, 1\}$ is attained for $d = 2$. When the covariance parameter r is parameter of interest, then the information about the parameter vector (θ_1, r) has the form $M = \frac{2d^2 \exp(-2rd)(1+\exp(-2rd))}{(1+\exp(-rd))(1-\exp(-2rd))^2}$, which attains its maximum for $d = 0$. We have $M_r = \frac{d^2 \exp(-2rd)(1+\exp(-2rd))}{(1-\exp(-2rd))^2}$, and $M_{\theta_1} = \frac{2}{1+\exp(-rd)}$ and the determinant of the Fisher information on (θ_1, r) is

$$M = \frac{2d^2 \exp(-2rd)(1+\exp(-2rd))}{(1+\exp(-rd))(1-\exp(-2rd))^2}.$$

Therein is also shown that if the only covariance parameter is of interest the maximal Fisher information is obtained for $d = 0$. To avoid such 'inconvenient' behavior it is suggested to decrease the non-diagonal elements by multiplying with a factor α, $0 < \alpha < 1$. By this we include a nugget effect (micro-scale variation effect) of the form

$$\gamma(d, r) = \begin{cases} 0, & \text{for } d = 0, \\ 1 - \alpha + \alpha(1 - \exp(-rd)), & \text{otherwise.} \end{cases}$$

Then we obtain

$$M_r = \frac{\alpha^2 d^2 \exp(-2dr)(\alpha^2 \exp(-2dr) + 1)}{(1 - \alpha^2 \exp(-2dr))^2}.$$

In [Stehlík, Rodríguez-Díaz, Müller and López-Fidalgo] is proved, that the distance d of the optimal design is an increasing function of nugget $1 - \alpha$.

For $n = 3$ we obtain

$$M_r(3)(d_{12}, d_{23}, r) = (d_{12}^2 \exp(-2rd_{12}) - 2\exp(-2r(d_{12}+d_{23}))d_{12}^2 - 2\exp(-2r(2d_{12}+d_{23}))d_{12}^2 +$$

$$+\exp(-2r(2d_{23}+d_{12}))d_{12}^2 - 2\exp(-2r(d_{12}+d_{23}))d_{23}^2 + d_{23}^2\exp(-2rd_{23}) + d_{23}^2\exp(-2r(2d_{12}+d_{23})) +$$
$$+d_{12}^2\exp(-4r(d_{12}+d_{23})) + d_{23}^2\exp(-4r(d_{12}+d_{23})) + d_{23}^2\exp(-4rd_{23}) + d_{12}^2\exp(-4rd_{12}) +$$
$$-2d_{23}^2\exp(-2r(2d_{23}+d_{12})))/(-1+\exp(-2rd_{12})+\exp(-2rd_{23})-\exp(-2r(d_{12}+d_{23})))^2.$$

Some numerical obstacles with collapsing of designs without nugget can exhibit here. For instance, when we are seeking for the maximum of $M_r(3)$ numerically, employing Maple 7 function `minimize`, we obtain that $\max_{0<d_{12}<1, 0<d_{23}<1} M_r(3)(d_{12}, d_{23}, r) = +\infty$, so the numerics could be misleading.

Let us have $n = 5$. The function $M_r(5)$ is very complex: the direct symbolic computation of $M_r(5)$ takes (Maple 7) $753.2sec$ and optimization of $M_r(5)$ by the routine `minimize` on the compact set $[0, 30]$ is not ready within 8 hours. Therefore we decided for a grid optimization. We have found (by R Version 1.8.1 and $r = 1$) that for $d_{12} = 18$, $d_{23} = 4.8$, $d_{34} = 7.2$ and $d_{45} = 0$ we obtain $M_r(5) = 2.671118e + 16$. For instance, the same function Maple 7 implementation gave for same input $M_r(5) = 597804719.6$ (numerical controversy). Employing of R did not led to a satisfactory results.

To get rid of these obstacles, following the observations of [Royle 02], we have employed an equidistant design $t = d_{12} = d_{23} = d_{34} = d_{45}$. [Royle 02] suggested to use so called space-filling designs, as an alternative to covariance-based design criteria. Such designs depend only on the spatial locations of design and candidate points.

Here we have

$$M_r(5) = 4t^2 \exp(-2rt)\frac{\exp(-2rt) + 1}{\exp(-4rt) - 2\exp(-2rt) + 1}$$

and $\lim_{t\to 0+} \det M_r(5) = 2/r^2$. Note that for an equidistant design we have $4M_r(2) = M_r(5)$, $3M_r(2) = M_r(4)$ and $2M_r(2) = M_r(3)$. Here one can construct the hypothesis, that for an equidistant design we have $k_n M_r(2) = M_r(n)$ for some $k_n > 0$, with $k_3 = 2, k_4 = 3, k_5 = 4$. We have checked numerically (on the grid $\{i/2\}_{i=1}^{99}$ for $r = 1, 5, 50$) that $k_{10} = 9, k_{15} = 14$ (and we have made some numerical checks justifying that $k_{20} = 19$) so our **conjecture** is

$$k_n = n - 1. \tag{1}$$

One can find a nice geometrical interpretation of this information behavior. Let us imagine that design points (vertexes) are connected with edges and constitute a simple tree (from Graph Theory, see e.g. [Foulds 92]), such that all vertices besides the beginning and ending have the degree two. Then adding another design point adds one edge. So the information relation $(n-1)M_r(2) = M_r(n)$ has direct Graph-ical representation and interpretation.

Conjecture 1 *Let us have an equidistant $(n+1)$-point design with distances $d_{12} = ... = d_{n,n+1} = t$. Then the r-information has under the (1) the form*

$$M_r(n+1) = nt^2 \exp(-2rt) \frac{\exp(-2rt) + 1}{\exp(-4rt) - 2\exp(-2rt) + 1}.$$

Conjecture 2 *The optimal equidistant $(n+1)$-point design for parameter r with distances $d_{12} = ... = d_{n,n+1} = t$ is under the (1) the collapsed one, i.e.* $\arg\max M_r(n+1) = \{0\}$. *This collapsing can be tuned by the nugget effect.*

The collapsing effect of an equidistant design is connected with the covariance parameter MLE behavior in Gaussian field with an exponential covariance structure. These properties are available e.g. from the analytical study of the log-likelihood limit for $rt \to 0$.

3 Exchange-type algorithm for computation of D-optimum designs

The design problem with correlated observations differs from the much studied classical design problem because the covariance matrix may here be non-diagonal, due to the correlations between different observations. This constitutes a major obstacle when trying directly to apply the numerical algorithms already existing in the optimum experimental design literature, which were constructed for experiments with uncorrelated observations. The technical point is that the algorithms make explicit use of the fact that, for uncorrelated observations, the Fisher information matrix can be expressed as the sum of the matrices of the individual points forming the design.

Generally, spatial design problems consist of selecting n-points in some m-dimensional space domain so as to optimize the value of an optimality criterion function Φ, in our case (D-optimality), $\Phi(M) = \det(M)$. Optimization of Φ may be difficult due to the complex nature of Φ (e.g. many local optima, complex derivation), and also the high dimensionality of the problem. Many have recognized these problems and have sought to develop optimization algorithms. One approach to optimization for design problems involves the use of "exchange" algorithms, whereby Φ is optimized marginally for elements of design space by successively exchanging points with those which produce improvements in Φ. The use of exchange algorithms is not new, some references are Brimkulov (1980) and Ucinski and Atkinson (2004). Exchange algorithms have found widespread use due to their efficiency and ability to handle very general problems. They are rather easily implemented for arbitrary shaped design regions and definition of candidate points and also for arbitrary design criteria. For very large problems, even use of exchange algorithms can become computationally prohibitive. Here large could mean a number of this including:

- the design criterion is itself computationally expensive to evaluate
- the candidate design space is large
- the design consists of a large number of points

3.1 Random start

The proposed algorithm is starting with the small number of random generated vectors, for instance independent, uniformly distributed. The following example illustrates the information of the random vectors. In our algorithm the most informative one is chosen to be an initial design.

Example.

Table Information of random initial designs

design	contains information
{0.8253546, 0.3216210, 0.4080212, 0.9146516, 0.2792454}	0.5644044
{0.53305164, 0.90556049, 0.53145502, 0.95335055, 0.01824999}	1.050907
{0.4492698, 0.5533527, 0.7828926, 0.9546576, 0.2927846}	0.6018488
{0.2432972, 0.7362840, 0.5309174, 0.3822175, 0.8843583}	0.5730896
{0.9508731, 0.7680115, 0.6334870, 0.8464303, 0.2846423}	0.6074925
{0.7533352, 0.6251179, 0.8555488, 0.7973968, 0.3569893}	0.3952017

3.2 Straightforward Algorithm

The algorithm proposed here is an exchange-type algorithm for computation of D-optimum designs inspired by [Brimkulov et al. 80] and [Ucinski and Atkinson 04].

Step 1. Select an initial design $x(0) = \{x_1(0), ..., x_n(0)\}$ obtained from the random start (see subsection Random start). Calculate the matrices $M_\theta(n)$, $M_r(n)$ and $M(n)$ and then the value of $\det M(n)$. If $\det M(n) = 0$, select a new initial design and repeat this step.

Step 2. Set $l = 0$.

Step 3. Determine $(i^\star, t^\star) = \arg\max\{\Delta(t_i, t) : (i, t) \in I \times T\}$, where

$$\Delta(t_i, t) = \frac{\det M(n)(x(l, x_i(l) \longleftrightarrow x)) - \det M(n)(x(l))}{\det M(n)(x(l))}.$$

Step 4. If $\Delta(t_i^\star, t^\star) \leq \delta$, where δ is some given positive tolerance, then Stop. Otherwise, set $x(l+1) = x(l, x_i(l) \longleftrightarrow x)$ and determine matrices corresponding to $x(l+1)$. Set $l \to l+1$ and go to Step 3.

4 Ucinski-Atkinson Example

The paper of Ucinski-Atkinson 04 is motivated by an example drawn from chemical kinetics. If the initial concentration of A is one and that of B and C are zero, the concentration of B is given by

$$\eta(t,\theta) = \frac{\theta_1}{\theta_1 - \theta_2}(exp(-\theta_2 t) - exp(-\theta_1 t)).$$

Let us start with two observations ($n = 2$). The information matrix $M_\theta(2)$ is a complex function. We obtain for the correlated observations almost the same result as Box and Lucas (1959) for uncorrelated case (and $n = 2$) and approximately, the optimal times are $t_1 = 1.23$ and $t_2 = 6.86$. To enhance the study of Ucinski and Atkinson, we consider also the information on correlation parameter. When r is also the parameter of interest (i.e. $M_r(2)$ is also included), we obtain the shift to a shorter times, e.g. for $\theta_1 = 0.7$, $\theta_2 = 0.2$, $r = 1$, and $t_1 = 1.23$ we obtain (approximative result) $t_2 \approx 2.98$. Ucinski and Atkinson have employed 1-dim search (`fminbnd`, Matlab v.6.5. Rel. 13) within the Brimkulov algorithm framework.

However, the function $M_\theta(5)$ is so complex, that we prefer to avoid the symbolic calculation of $\det M_\theta(5)$. By exchange-type optimization the result of Ucinski and Atkinson is "justified" for $n = 5$ and $r = 1$. Using of the direct algorithm on the distances $(d_{12}, d_{23}, d_{34}, d_{45})$, grid $\{1/2, 1, ..., 9/2\}^4$ and $\theta_1 = 0.7$, $\theta_2 = 0.2$, $t_1 = 1$ leads to the designs given in Table 1. As we can see, the effect of correlation parameter information is contraction of the distances $d_{12}, d_{23}, d_{34}, d_{45}$. Actually, there are two outstanding properties: The greater the correlation, the more the optimal observations are spread over the region (as pointed out formerly by Ucinski and Atkinson) and this effect becomes less significant when r is also the interest parameter. When we run the finite grid optimization with the net $\{1/2, 1, ..., 9/2\}^5$, we obtain the optimum for $(t_1, d_{12}, d_{23}, d_{34}, d_{45}) = (1, 1, 7/2, 5/2, 5/2)$, which justifies the previous results.

Table 1:n=5.

r	(θ_1, θ_2, r)	(θ_1, θ_2)	r
1	$\{1, 1.5, 6, 6.5, 7.5\}$	$\{1, 2, 5.5, 8, 10.5\}$	$\{t_1, t_1 + 0.5, t_1 + 1, t_1 + 1.5, t_1 + 2\}$
5	$\{1, 1.5, 6, 6.5, 7\}$	$\{1, 1.5, 6, 7, 8\}$	$\{t_1, t_1 + 0.5, t_1 + 1, t_1 + 1.5, t_1 + 2\}$
50	$\{1, 1.5, 6, 6.5, 7\}$	$\{1, 1.5, 6, 6.5, 7\}$	$\{t_1, t_1 + 0.5, t_1 + 1, t_1 + 1.5, t_1 + 2\}$

For a finer net $\{1/19, 2/19, ..., 1\}^4$, we obtain (putting $r = 1, 5, 50$) the same collapsing design $d_{12} = d_{23} = d_{34} = d_{45} = 1/19$, this is some justifying of the $M_r(5)$ contracting effect in our setup. For $n = 10$ we have the same equidistant optimal design on the fine grid $\{1/2, 1, 3/2\}^9$.

The D_S**-optimality** for both correlation parameters r and trend parameters θ is identical with D-optimality for r (i.e. the argument maxima of $\det M_r(n)$) and D-optimality for θ (i.e. the argument maxima of $\det M_\theta(n)$).

Using of the direct algorithm on the distances $(d_{12}, ..., d_{9,10})$, grid $\{1/2, 1, 3/2\}^9$ and $\theta_1 = 0.7$, $\theta_2 = 0.2$, $t_1 = 1$ leads to the designs given in Table 2.

Table 2:n=10.

r	(θ_1, θ_2, r)	(θ_1, θ_2)	r
1	$d = 1/2$	$\{1, 1.5, 2.5, 4, 5.5, 7, 8.5, 10, 11.5, 13\}$	$d = 1/2$
5	$\{1, 1.5, 3, 3.5, 5, 5.5, 7, 7.5, 9, 9.5\}$	$\{1, 1.5, 2, 2.5, 4, 5.5, 6.5, 7.5, 8.5, 9.5\}$	$d = 1/2$
50	$d = 1/2$	$\{1, 1.5, 2, 2.5, 4, 5.5, 6.5, 7, 7.5, 8\}$	$d = 1/2$

Using of the direct algorithm on the distances $(d_{12}, ..., d_{14,15})$, grid $\{1/2, 1\}^{14}$ and $\theta_1 = 0.7, \theta_2 = 0.2$, $t_1 = 1$ leads to the designs given in Table 3. When the only r is the parameter of interest, then $d = 1/2$ (equidistant design).

Table 3:n=15

r	(θ_1, θ_2, r)	(θ_1, θ_2)
1	$1, 1.5, 2, 2.5, 3, 3.5, 4, 4.5, 5, 5.5, 6, 7, 7.5, 8, 8.5$	$1, 1.5, 2, 2.5, 3, 3.5, 4, 4.5, 5, 5.5, 6, 7, 7.5, 8, 8.5$
5	$1, 1.5, 2.5, 3, 3.5, 4, 4.5, 5, 5.5, 6, 6.5, 7, 7.5, 8.5, 9$	$1, 1.5, 2.5, 3, 3.5, 4, 4.5, 5, 5.5, 6, 6.5, 7, 7.5, 8.5, 9$
50	$d = 1/2$	$d = 1/2$

The relative efficiency of the Brimkulov algorithm in Ucinski and Atkinson example can be tuned by δ and grid structure. For instance, when r is the parameter of interest, for the stopping rule with $\delta = 0.001$, $n = 15$, starting design $d_{12} = 9.5, d_{i+1,i+2} = 1/2$, and grid $\{1/2, 1, .., 10.5\}$ we have the convergence of Brimkulov algorithm to the finite grid optimum, an equidistant design with $d = 1/2$. However, at simple grid $\{1/2, 1\}$ we have no improvement for, approximately, $\delta > 0.2.10^{-6}$.

5 Conclusions

In the present paper we have studied some properties of exchange algorithms employed with D-optimal designing under the presence of correlation. Also in higher dimensions the exponential structure without nugget leads to collapsing of the design when also parameters of covariance structure are of interest, as have been also observed by [Zhu and Stein 04] among others. However, if only trend parameters are of interest, the designs covering uniformly the whole design space are very efficient.

Acknowledgement

Research was partly supported by FWF-grant (Austrian grant agency) P16809-N2, by the VEGA grant (Slovak grant agency) No 1/3016/06 and by a Forschungstipendium of Vice-Rektor for research of Wirtschaftsuniversität Wien. I am thankful for fruitful comments of Werner Müller during the preparation of paper.

References

[Abramowitz and Stegun 65] Abramowitz, M. and Stegun, I. A. (1965), Handbook of Mathematical Functions, New York, Dover.

[Abt and Welch 98] Abt M. and Welch W.J. (1998). *Fisher information and maximum-likelihood estimation of covariance parameters in Gaussian stochastic processes*, The Canadian Journal of Statistics, Vol. 26, No. 1, 127-137.

[Box and Lucas 59] Box G.E.P. and Lucas H.L., Design of experiments in nonlinear situations, *Biometrika 46*, 77-90.

[Brimkulov et al. 80] Brimkulov, U.N., Krug, G.K. and Savanov, V.L. (1980). Numerical construction of exact experimental designs when the measurements are correlated (in Russian). *Zavodskaya Laboratoria (Industrial Laboratory)*, 36:435–442.

[Foulds 92] Foulds L.R. (1992) Graph Theory Applications, Springer Verlag, New York, Inc., p. 17-19.

[Handcock and Wallis 94] Handcock, M.S. and Wallis, J.R.(1994) An Approach to Statistical Spatial-Temporal Modeling of Meteorological Fields: Journal of the American Statistical Association, Vol. 89, pp. 368-378.

[Horst and Tuy 96] Horst, R. and Tuy, H. (1996) Global Optimization - Deterministic Approaches, Springer.

[Müller and Stehlík 04] Müller W. G. and Stehlík M. (2004), *An example of D-optimal designs in the case of correlated errors*, Proceedings in Computational Statistics, Edited by J. Antoch, Physica-Verlag, 1543-1550.

[Pázman 04] Pázman A. (2004). *Correlated optimum design with parametrized covariance function: Justification of the use of the Fisher information matrix and of the method of virtual noise*, Research Report Series of the Department of Statistics and Mathematics, Wirtschaftsuniversität Wien, Austria.

[Royle 02] Royle J. A. (2002). Exchange algorithms for constructing large spatial designs, Journal of Statistical Planning and Inference 100, p. 121-134.

[Stehlík 04] Stehlík M., (2004), *Some Properties of D-optimal Designs for Random Fields with Different Variograms*, Report #4 of the Research Report Series of the Department of Statistics and Mathematics, Wirtschaftsuniversität Wien, Austria.

[Stehlík, Rodríguez-Díaz, Müller and López-Fidalgo] Stehlík M., Rodríguez-Díaz J.M., Müller W.G. and López-Fidalgo J. Optimal allocation of bioasseys in the case of parametrized covariance functions: an application in lung's retention of radioactive particles, *TEST* (in press).

[Stehlík 05] Stehlík M., (2005), Covariance related properties of D-optimal correlated designs, Proceedings of the 5th St. Petersburg Workshop on Simulation, St. Petersburg, Russia, June 26-July 2, Edited by S.M. Ermakov, V.B. Melas and A.N. Pepelyshev, 645-652, 2005.

[Ucinski and Atkinson 04] Ucinski D. and Atkinson A. C. (2004) "Experimental Design for Time-Dependent Models with Correlated Observations", Studies in Nonlinear Dynamics & Econometrics: Vol. 8: No. 2, Article 13. http://www.bepress.com/snde/vol8/iss2/art13

[Zhang 04] Zhang, H. (2004). Inconsistent estimation and asymptotically equivalent interpolations in model-based geostatistics. JASA, 99, 250-61.

[Zhang and Zimmerman 05] Zhang H. and Zimmerman D. (2005), Towards reconciling two asymptotic framework in spatial statistics, Biometrika, 92, 4, pp. 921-936.

[Zhu and Stein 04] Zhu, Z. and Stein, ML (2004), Spatial sampling design for parameter estimation of the covariance function, Journal of Statistical Planning and Inference, in press.

[Zhu and Zhang 05] Zhu, Z. and Zhang, H (2005), Spatial sampling design under the Infill Asymptotic Framework, Environmetrics, in press.